

1600

## RAW SEQUENCE LISTING

DATE: 09/07/2001

PATENT APPLICATION: US/09/597,796A

TIME: 13:55:04

Input Set : A:\-90-7pc.app

Output Set: N:\CRF3\09072001\I597796A.raw

3 <110> APPLICANT: Skeiky, Yasir  
 4 Reed, Steven  
 5 Alderson, Mark  
 6 Corixa Corporation  
 8 <120> TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis  
 10 <130> FILE REFERENCE: 014058-009070PC  
 C--> 12 <140> CURRENT APPLICATION NUMBER: US/09/597,796A  
 C--> 13 <141> CURRENT FILING DATE: 2000-06-20  
 15 <150> PRIOR APPLICATION NUMBER: US 09/597,796  
 16 <151> PRIOR FILING DATE: 2000-06-20  
 18 <150> PRIOR APPLICATION NUMBER: US 60/265,737  
 19 <151> PRIOR FILING DATE: 2001-02-01  
 21 <160> NUMBER OF SEQ ID NOS: 49  
 23 <170> SOFTWARE: PatentIn Ver. 2.1  
 25 <210> SEQ ID NO: 1  
 26 <211> LENGTH: 1872  
 27 <212> TYPE: DNA  
 28 <213> ORGANISM: Mycobacterium tuberculosis  
 30 <220> FEATURE:  
 31 <223> OTHER INFORMATION: MTB32A (Ra35FL)  
 33 <220> FEATURE:  
 34 <221> NAME/KEY: modified\_base  
 35 <222> LOCATION: (1)..(1872)  
 36 <223> OTHER INFORMATION: n = g, a, c or t  
 38 <400> SEQUENCE: 1  
 39 gactacgttg gtgtagaaaa atcctgcccgc cgggaccctt aaggctggga caatttctga 60  
 40 tagctacccc gacacaggag gttacgggat gagcaattcg cgccgcccgt cactcagggtg 120  
 41 gtcatggttg ctgagcgtgc tggctgcccgt cgggctgggc ctggccacgg cgccggccca 180  
 42 ggcggccccc cgggccttgt cgcaggaccg gttcgcgcac ttccccgcgc tgcccctcga 240  
 43 cccgtcccgc atggtcgccc aagtggcgcc acagggtggtc aacatcaaca ccaaactggg 300  
 44 ctacaacaac gccgtgggcg ccgggaccgg catcgtcatc gatcccaacg gtgtcgtgct 360  
 45 gaccaacaac cacgtgatcg cgggcgccac cgacatcaat gcgttcagcg tcggctccgg 420  
 46 ccaaacctac ggcgtcgatg tggtcgggta tgaccgcacc caggatgtcg cgggtgctgca 480  
 47 gctgcgcggt gccggtggcc tgccgtcggc ggcgatcggg ggcggcgtcg cggttggtga 540  
 48 gcccgctgct gcgatgggca acagcgggtg gcaggggcga acgccccgtg cggtgccctgg 600  
 49 cagggtggtc gcgctcggcc aaaccgtgca ggcgtcggat tcgctgaccg gtgcccgaaga 660  
 50 gacattgaac ggggtgatcc agttcgatgc cgcaatccag cccggtgatt cgggcccggcc 720  
 51 cgctcgtaac ggcctaggac aggtggtcgg tatgaacacg gccgcgtccg ataacttcca 780  
 52 gctgtcccag ggtgggcagg gattcgccat tccgatcggg caggcgatgg cgatcgcggg 840  
 53 ccaaataccg tcgggtgggg ggtcaccac cggtcatatc gggcctaccg ccttcctcgg 900  
 54 cttgggtggt gtcgacaaca acggcaacgg cgacagagtc caacgcgtgg tcggaagcgc 960  
 55 tccggcggca agtctcggca tctccaccgg cgacgtgatc accgcggtcg acggcgctcc 1020  
 56 gatcaactcg gccaccgcga tggcggacgc gcttaacggg catcatcccg gtgacgtcat 1080  
 57 ctcggtgaac tggcaaacca agtcgggcgg cagcgtaca gggaacgtga cattggccga 1140  
 58 gggacccccg gcctgatttg tcgcggatag caccgcggcg ccggccaatt ggattggcgc 1200  
 59 cagccgtgat tgccgcgtga gccccgagt tccgtctccc gtgcgcgtgg cattgtggaa 1260  
 60 gcaatgaacg aggcagaaca cagcgttgag caccctcccg tgcagggcag ttacgtcgaa 1320

ENTERED

See page 5

## RAW SEQUENCE LISTING

DATE: 09/07/2001

PATENT APPLICATION: US/09/597,796A

TIME: 13:55:04

Input Set : A:\-90-7pc.app

Output Set: N:\CRF3\09072001\I597796A.raw

```

61 ggcgggtgtgg tcgagcatcc ggatgccaaag gacttcggca gcgcgcgcgc cctgcccgcc 1380
62 gatccgacct ggtttaagca cgccgtcttc tacgaggtgc tggtcggggc gttcttcgac 1440
W--> 63 gccagcgcgg acggttccgn cgatctgcgt ggactcatcg atcgccctga ctacctgcag 1500
64 tggcttggca tcgactgcat ctgttgccgc cgttcctacg actcaccgct gcgcgacggc 1560
65 ggttacgaca ttcgcgactt ctacaagggtg ctgcccgaat tcggcaccgt cgacgatttc 1620
66 gtcgccttg tcgacaccgc tcaccggcga ggtatccgca tcatcaccga cctggtgatg 1680
67 aatcacacct cggagtcgca cccctggttt caggagtccc gccgcgaccc agacggaccg 1740
68 tacggtgact attacgtgtg gacgcacacc agcgcgcgct acaccgacgc ccggtatcatc 1800
W--> 69 ttcgtcgaca ccgaagagtc gaactggtca ttcgatcctg tccgcgcgaca gttinctactg 1860
70 gcaccgattc tt 1872
73 <210> SEQ ID NO: 2
74 <211> LENGTH: 355
75 <212> TYPE: PRT
76 <213> ORGANISM: Mycobacterium tuberculosis
78 <220> FEATURE:
79 <223> OTHER INFORMATION: MTB32A (Ra35FL)
81 <400> SEQUENCE: 2
82 Met Ser Asn Ser Arg Arg Arg Ser Leu Arg Trp Ser Trp Leu Leu Ser
83 1 5 10 15
85 Val Leu Ala Ala Val Gly Leu Gly Leu Ala Thr Ala Pro Ala Gln Ala
86 20 25 30
88 Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
89 35 40 45
91 Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Ala Pro Gln Val Val
92 50 55 60
94 Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
95 65 70 75 80
97 Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
98 85 90 95
100 Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln
101 100 105 110
103 Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
104 115 120 125
106 Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
107 130 135 140
109 Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
110 145 150 155 160
112 Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
113 165 170 175
115 Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr
116 180 185 190
118 Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser
119 195 200 205
121 Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
122 210 215 220
124 Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe Ala
125 225 230 235 240
127 Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser Gly
128 245 250 255

```

## RAW SEQUENCE LISTING

DATE: 09/07/2001

PATENT APPLICATION: US/09/597,796A

TIME: 13:55:04

Input Set : A:\-90-7pc.app

Output Set: N:\CRF3\09072001\I597796A.raw

```

130 Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly Leu
131          260          265          270
133 Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val Val
134          275          280          285
136 Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val Ile
137          290          295          300
139 Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala Asp
140 305          310          315          320
142 Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Asn Trp Gln
143          325          330          335
145 Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu Gly
146          340          345          350
148 Pro Pro Ala
149          355
152 <210> SEQ ID NO: 3
153 <211> LENGTH: 1002
154 <212> TYPE: DNA
155 <213> ORGANISM: Mycobacterium tuberculosis
157 <220> FEATURE:
158 <223> OTHER INFORMATION: MTB32A (Ra35 mature)
160 <400> SEQUENCE: 3
161 catatgcatac accatcacca tcacgccccg ccggccttgt cgcaggaccg gttcgcgcgac 60
162 ttccccgcgc tgccccctcga cccgtccgcg atggctcgccc aagtggggcc acagggtggtc 120
163 aacatcaaca ccaaactggg ctacaacaac gccgtgggcg ccgggaccgg catcgtcatac 180
164 gatcccaacg gtgtcgtgct gaccaacaac cacgtgatcg cgggcgccac cgacatcaat 240
165 gcgttcagcg tcggctccgg ccaaacctac ggcgtcgatg tggtcgggta tgaccgcacc 300
166 caggatgtcg cgggtgctgca gctgcgcggt gccggtggcc tgccgtcggc ggcgatcggg 360
167 ggcggcgctg cgggttggtga gcccgtcgtc gcgatgggca acagcgggtg gcagggcgga 420
168 acgccccgtg cgggtgcctgg cagggtggtc gcgctcggcc aaaccgtgca ggcgtcggat 480
169 tcgctgaccg gtgcccgaaga gacattgaac gggttgatcc agttcgatgc cgcgatccag 540
170 cccggtgagg cgggcggggc cgtcgtcaac ggcctaggac aggtgggtcg tatgaacacg 600
171 gccgcgtccg ataacttcca gctgtcccag ggtgggcagg gattcgccat tccgatcggg 660
172 caggcgatgg cgatcgcggg ccagatccga tcgggtgggg ggtcaccac cgttcataat 720
173 gggcctaccg ctttctctcg cttgggtgtt gtcgacaaca acggcaacg cgcacgagtc 780
174 caacgcgtgg tcgggagcgc tccggcgcca agtctcgga tctccaccg cgacgtgatc 840
175 accgcggtcg acggcgctcc gatcaactcg gccaccgca tggcggacgc gcttaacggg 900
176 catcatcccg gtgacgtcat ctcggtgacc tggcaaacca agtcgggcgg cacgcgtaca 960
177 gggaacgtga cattggccga gggacccccg gcctgagaat tc 1002
180 <210> SEQ ID NO: 4
181 <211> LENGTH: 330
182 <212> TYPE: PRT
183 <213> ORGANISM: Mycobacterium tuberculosis
185 <220> FEATURE:
186 <223> OTHER INFORMATION: MTB32A (Ra35 mature)
188 <400> SEQUENCE: 4
189 Met His His His His His His Ala Pro Pro Ala Leu Ser Gln Asp Arg
190 1          5          10          15
192 Phe Ala Asp Phe Pro Ala Leu Pro Leu Asp Pro Ser Ala Met Val Ala
193          20          25          30

```

## RAW SEQUENCE LISTING

DATE: 09/07/2001

PATENT APPLICATION: US/09/597,796A

TIME: 13:55:04

Input Set : A:\-90-7pc.app

Output Set: N:\CRF3\09072001\I597796A.raw

```

195 Gln Val Gly Pro Gln Val Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn
196          35          40          45
198 Asn Ala Val Gly Ala Gly Thr Gly Ile Val Ile Asp Pro Asn Gly Val
199          50          55          60
201 Val Leu Thr Asn Asn His Val Ile Ala Gly Ala Thr Asp Ile Asn Ala
202 65          70          75          80
204 Phe Ser Val Gly Ser Gly Gln Thr Tyr Gly Val Asp Val Val Gly Tyr
205          85          90          95
207 Asp Arg Thr Gln Asp Val Ala Val Leu Gln Leu Arg Gly Ala Gly Gly
208          100          105          110
210 Leu Pro Ser Ala Ala Ile Gly Gly Gly Val Ala Val Gly Glu Pro Val
211          115          120          125
213 Val Ala Met Gly Asn Ser Gly Gly Gln Gly Gly Thr Pro Arg Ala Val
214          130          135          140
216 Pro Gly Arg Val Val Ala Leu Gly Gln Thr Val Gln Ala Ser Asp Ser
217 145          150          155          160
219 Leu Thr Gly Ala Glu Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala
220          165          170          175
222 Ala Ile Gln Pro Gly Asp Ser Gly Gly Pro Val Val Asn Gly Leu Gly
223          180          185          190
225 Gln Val Val Gly Met Asn Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser
226          195          200          205
228 Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile
229          210          215          220
231 Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly
232 225          230          235          240
234 Pro Thr Ala Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly
235          245          250          255
237 Ala Arg Val Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly
238          260          265          270
240 Ile Ser Thr Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn
241          275          280          285
243 Ser Ala Thr Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp
244          290          295          300
246 Val Ile Ser Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly
247 305          310          315          320
249 Asn Val Thr Leu Ala Glu Gly Pro Pro Ala
250          325          330
253 <210> SEQ ID NO: 5
254 <211> LENGTH: 1002
255 <212> TYPE: DNA
256 <213> ORGANISM: Artificial Sequence
258 <220> FEATURE:
259 <223> OTHER INFORMATION: Description of Artificial Sequence:Ra35FLMutSA
261 <400> SEQUENCE: 5
262 catatgcattc accatcacca tcacgccccg ccggccttgt cgcaggaccg gttcgccgac 60
263 ttccccgcgc tgccccctcga cccgtccgcg atggtcgccc aagtggggcc acaggtggtc 120
264 aacatcaaca ccaaaactggg ctacaacaac gccgtgggcg ccgggaccgg catcgtcatt 180
265 gatcccaacg gtgtcgtgct gaccaacaac cacgtgatcg cgggcgccac cgacatcaat 240

```

## RAW SEQUENCE LISTING

DATE: 09/07/2001

PATENT APPLICATION: US/09/597,796A

TIME: 13:55:04

Input Set : A:\-90-7pc.app

Output Set: N:\CRF3\09072001\I597796A.raw

```

266 gcgttcagcg tcggctccgg ccaaacctac ggcgtcgatg tggtcgggta tgaccgcacc 300
267 caggatgtcg cggtgctgca gctgcgcggt gccggtggcc tgccgtcggc ggcgatcggg 360
268 ggcggcgctcg cggttggtga gcccgctcgtc gcgatgggca acagcgggtg gcagggcgga 420
269 acgccccgtg cggtgcctgg caggggtggc gcgctcggcc aaaccgtgca ggcgtcggat 480
270 tcgctgaccg gtgccgaaga gacattgaac gggttgatcc agttcgatgc cgcgatccag 540
271 cccggtgatg cgggcgggcc cgtcgtcaac ggcctaggac aggtggtcgg tatgaacacg 600
272 gccgcgtccg ataacttcca gctgtcccag ggtgggcagg gattcgccat tccgatcggg 660
273 caggcgatgg cgatcgcggg ccagatccga tcgggtgggg ggtcaccac cgttcatatc 720
274 gggcctaccg ccttctcggg cttgggtggt gtcgacaaca acggcaacgg cgcacgagtc 780
275 caacgcgtgg tcgggagcgc tcgggcggca agtctcggca tctccaccgg cgacgtgac 840
276 accgcggtcg acggcgctcc gatcaactcg gccaccgcga tggcggagcg gcttaacggg 900
277 catcatcccg gtgacgtcat ctcgggtgacc tggcaaacca agtcgggcgg caccggtaca 960
278 gggaacgtga cattggccga gggacccccg gcctgagaat tc 1002
281 <210> SEQ ID NO: 6
282 <211> LENGTH: 330
283 <212> TYPE: PRT
284 <213> ORGANISM: Artificial Sequence
286 <220> FEATURE:
287 <223> OTHER INFORMATION: Description of Artificial Sequence:Ra35FLMutSA
289 <400> SEQUENCE: 6
290 Met His His His His His His Ala Pro Pro Ala Leu Ser Gln Asp Arg
291 1 5 10 15
293 Phe Ala Asp Phe Pro Ala Leu Pro Leu Asp Pro Ser Ala Met Val Ala
294 20 25 30
296 Gln Val Gly Pro Gln Val Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn
297 35 40 45
299 Asn Ala Val Gly Ala Gly Thr Gly Ile Val Ile Asp Pro Asn Gly Val
300 50 55 60
302 Val Leu Thr Asn Asn His Val Ile Ala Gly Ala Thr Asp Ile Asn Ala
303 65 70 75 80
305 Phe Ser Val Gly Ser Gly Gln Thr Tyr Gly Val Asp Val Val Gly Tyr
306 85 90 95
308 Asp Arg Thr Gln Asp Val Ala Val Leu Gln Leu Arg Gly Ala Gly Gly
309 100 105 110
311 Leu Pro Ser Ala Ala Ile Gly Gly Gly Val Ala Val Gly Glu Pro Val
312 115 120 125
314 Val Ala Met Gly Asn Ser Gly Gly Gln Gly Gly Thr Pro Arg Ala Val
315 130 135 140
317 Pro Gly Arg Val Val Ala Leu Gly Gln Thr Val Gln Ala Ser Asp Ser
318 145 150 155 160
320 Leu Thr Gly Ala Glu Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala
321 165 170 175
323 Ala Ile Gln Pro Gly Asp Ala Gly Gly Pro Val Val Asn Gly Leu Gly
324 180 185 190
326 Gln Val Val Gly Met Asn Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser
327 195 200 205
329 Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile
330 210 215 220
332 Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly

```

FYI:

Use of n and / or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to ensure a corresponding explanation is present in the <220> to <223> fields of each sequence using n or Xaa.

## VERIFICATION SUMMARY

DATE: 09/07/2001

PATENT APPLICATION: US/09/597,796A

TIME: 13:55:05

Input Set : A:\-90-7pc.app

Output Set: N:\CRF3\09072001\I597796A.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:63 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:69 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:506 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11  
L:570 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12  
L:755 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15  
L:939 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15  
L:941 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:15  
L:950 M:258 W: Mandatory Feature missing, <220> FEATURE:  
L:1464 M:258 W: Mandatory Feature missing, <220> FEATURE:  
L:1704 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25  
L:1714 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25  
L:1720 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25  
L:1721 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25  
L:1743 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26  
L:2164 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36  
L:2442 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42  
L:2444 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42  
L:2447 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42  
L:2448 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42  
L:2474 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43  
L:2486 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43  
L:2516 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43  
L:2657 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:46  
L:2666 M:258 W: Mandatory Feature missing, <220> FEATURE:  
L:2930 M:258 W: Mandatory Feature missing, <220> FEATURE: